

## [SPECIFICATION]

## Method for Gene Analysis

5                   Background of the Inventiona           1. Field of the Invention

          The present invention relates to a method for gene  
analysis by hybridization, and more specifically, it  
10 relates to a novel method for gene analysis preferably  
utilized for gene analysis by hybridization using a DNA  
chip and the like, which method can be used for the  
analysis of nucleic acids, such as nucleotide sequence  
determination of nucleic acids, gene diagnosis of  
15 infectious diseases or genetic diseases, and monitoring  
of genome DNA expression.

a           2. Description of the Related Art

          The analysis of genes by a hybridization technique  
20 utilizing a substrate (DNA chip or DNA array) having  
immobilized probe nucleic acids is widely utilized, for  
example, for nucleotide sequence determination, gene  
diagnosis of infectious and genetic diseases, monitoring  
expression of genome gene and so forth. For instance,  
25 SBH [Sequencing By Hybridization, R. Drmanac et al.,  
Science, 260, 1649 (1993)], i.e., nucleotide sequencing  
via hybridization, is expected to be put into practical

use as a high-speed and low cost method. Further, the method for detecting mutations contained in genes using DNA chips [J.G. Hacia et al., Nature Genetics, 14,411-447(1996)] and the method of monitoring gene expression patterns using DNA chips [M. Schena et al., Science, 270,467-470(1995)] are drawing attention as methods enabling quick analysis of significant amounts of gene expression.

For these analyses of genes, an article called DNA chip or DNA array (DNA microarray or DNA macroarray) is utilized, which comprises nucleic acids such as DNA and/or RNA immobilized on a substrate.

As a substrate on which DNA is immobilized, membranes made of resin such as nylon membranes and polypropylene membranes, nitrocellulose membranes, glass plates or silicon plates are utilized. When detection of hybridization is performed by not using radioisotopes, but using, for example, fluorescent substances, it is preferable to utilize glass plates or silicon plates which contain no fluorescent substance.

However, when hybridization is performed on a DNA chip or DNA array (DNA microarray or DNA macroarray), the hybridization velocity is the major factor affecting the speed of gene analysis. The hybridization time required for gene analysis utilizing usual DNA chips or DNA arrays (DNA microarrays or DNA macroarrays) is 1 to 5 hours for gene sequence analysis [J.G. Hacia et al.,

Nature Genetics, 14, 441-447 (1996)], or 6 to 12 hours for gene sequence analysis [M. Shena et al., Proc. Natl. Acad. Sci. USA, Vol. 93, pp. 10614-10619 (1996)], and this has been a major problem in realizing gene analysis  
5 of higher speed utilizing the aforementioned method.

#### Summary of the Invention

The object of the present invention is to increase the speed of gene analysis by hybridization utilizing a  
10 probe nucleic acid.

The inventors of the present invention earnestly studied in order to achieve the aforementioned object. As a result, it was found that gene analysis speed can be increased by performing hybridization on a DNA chip or DNA array (DNA microarray or DNA macroarray), which  
15 consists of a substrate having an immobilized probe nucleic acid, in the presence of a double-stranded DNA-binding protein derived from a thermophilic bacterium. That is, it is considered that a double-stranded DNA-  
20 binding protein was bound to a double-stranded DNA in a DNA hybridization system in an equilibrated state to forward the reaction in a single direction (complementary double-stranded DNA forming direction), and that the heat resistance of the protein enabled the  
25 reaction at high temperatures and thereby realized higher speed and high-throughput of the gene analysis. The present invention has been accomplished on the basis

of these findings.

That is, the present invention provides a method for gene analysis comprising the step of detecting hybridization between a probe nucleic acid and a sample nucleic acid containing a target sequence that has a sequence complementary to that of the probe nucleic acid, wherein either the probe nucleic acid or the sample nucleic acid is immobilized on a substrate, at least one of the probe nucleic acid and the sample nucleic acid is DNA, and the hybridization is caused in the presence of a double-stranded DNA-binding protein.

The present invention also provides the aforementioned method for gene analysis wherein the sample nucleic acid is DNA.

The present invention further provides the aforementioned method for gene analysis wherein the double-stranded DNA-binding protein is derived from a hyperthermophilic bacterium.

The present invention further provides the aforementioned method for gene analysis wherein the double-stranded DNA-binding protein is derived from an archaeobacterium.

The present invention further provides the aforementioned method for gene analysis wherein the double-stranded DNA-binding protein is derived from a bacterium belonging to the genus *Sulfolobus*.

The present invention further provides the

aforementioned method for gene analysis wherein the double-stranded DNA-binding protein is derived from *Sulfolobus solfataricus*.

The present invention further provides the  
5 aforementioned method for gene analysis wherein the double-stranded DNA-binding protein is Sso7d protein derived from *Sulfolobus solfataricus*.

The present invention further provides the  
aforementioned method for gene analysis wherein the  
10 double-stranded DNA-binding protein is a protein having homology of 75% or more to the protein represented by the amino acid sequence of SEQ ID NO: 9.

The present invention further provides the  
aforementioned method for gene analysis wherein the  
15 sample nucleic acid is labeled.

The present invention further provides the  
aforementioned method for gene analysis wherein amount  
of the sample nucleic acid containing the target  
sequence is analyzed based on intensity of hybridization  
20 signal.

The present invention further provides the  
aforementioned method for gene analysis wherein  
detecting hybridization is performed by using a  
plurality of probe nucleic acids and then polymorphism  
25 in the target sequence is detected based on the result  
of detection of hybridization.

The present invention further provides the

aforementioned method for gene analysis wherein detecting hybridization is performed by using a plurality of probe nucleic acids and then nucleotide sequence of the sample nucleic acid is determined based on the result of detection of hybridization.

The present invention further provides a test kit for detection of hybridization between a probe nucleic acid and a sample nucleic acid containing a target sequence that has a sequence complementary to that of the probe nucleic acid, which comprises at least a double-stranded DNA-binding protein.

[DETAILED] DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention will be explained in detail hereafter.

In the present invention, the term "double-stranded DNA-binding protein" refers to a protein which binds to chromosome of eucaryote or that of prokaryote strongly and concerns retention of higher-order structure of chromosome. That is, it comprises a protein having function to stabilize a complementary double-stranded DNA.

In the present invention, the term "sample nucleic acid" refers to a nucleic acid which is a subject of analysis such as nucleotide sequence determination or expression analysis, and it may be either DNA or RNA.

In the present invention, the term "probe nucleic

acid" refers to a nucleic acid which is utilized for detection of a target gene in a sample nucleic acid, and which may be either DNA or RNA. Examples of the probe nucleic acid include a probe containing an  
5 oligonucleotide comprising usual base of nucleic acid that is A (adenine), T (thymine), G (guanine), C (cytosine) and U(uracil), and a probe consisting of a DNA fragment amplified by PCR and having a length of approximately 50-2,000 nucleotides. The length of the  
10 probe nucleic acid is not particularly limited so long as it is a length hybridizable with the sample nucleic acid. Those having a length of 6-90 nucleotides, preferably 8-30 nucleotides, are usually used. However, nucleic acids of either longer or shorter than these  
15 lengths may also be used.

As the probe nucleic acid, there may further be used a probe nucleic acid containing an oligonucleotide which comprises a modified nucleotide, for example, hypoxanthines such as inosine (Japanese Patent Laid-open  
20 (Kokai) No. 8-70900(U.S. Patent No. 5,738,993)), 5-nitroindole, 3-nitropyrrole (Japanese Patent Laid-open (Kokai) No. 10-262675), 2-aminoadenine, 5-(1-propynyl)uracil (Tetrahedron Letters, Vol.33, p. 5307-5310) and so forth.

25 In the probe nucleic acid, as described in Japanese Patent Laid-open (Kokai) No. 8-70900 (U.S. Patent No. 5,738,993), a non-specific region may be

ligated to at least one of the ends of the sequence  
region substantially complementary to the target  
sequence in the sample nucleic acid. Ligation of such a  
non-specific region enables increase of hybridization  
5 sensitivity as well as facilitates differentiation  
between a complementary hybrid and a hybrid having a  
mismatch, in particular, end mismatch.

The non-specific region consists of at least one  
nucleotide which has a base that can form a base pair  
10 with a nucleotide constituting normal nucleic acids, but  
is different from those of the nucleotides constituting  
normal nucleic acids, or an oligomer thereof. Such a  
base is preferably one that can associate with any of  
bases constituting normal nucleic acid equally, and  
15 strength of such association is preferably weaker than  
that of the base pairs constituting a specific pairing.  
As a specific example of such a base, hypoxanthine, 5-  
nitroindole, 3-nitropyrrole and so forth can be  
mentioned. As a specific example of such a nucleotide,  
20 there can be mentioned deoxyinosine, which is a  
deoxyribonucleotide having hypoxanthine as a base.

The number of nucleotides or oligomers thereof  
constituting the non-specific region is not particularly  
limited. However, it is preferably 2-20, more  
25 preferably 2-8. The location for ligation of the non-  
specific region to the specific region is not  
particularly limited as well, and it may be either of



the 5' end and 3' end of the specific region. Both of the 5' end and 3' end of the specific region may be each ligated with a non-specific region. Among these cases, the latter case is particularly preferred. While the ratio of lengths of the specific region and the non-specific region may vary depending on the length of the specific region or the GC content, the length of the specific region is preferably equal to or longer than that of the non-specific region.

10 By providing a non-specific region in the probe nucleic acid as described above, the gene analysis can be performed with higher precision.

The aforementioned probe nucleic acid and the sample nucleic acid can easily be synthesized by using a commercially available DNA synthesizer.

15 The probe nucleic acid is utilized to detect a sample nucleic acid comprising a nucleotide sequence complementary to that of the probe nucleic acid. This complementary sequence to the probe nucleic acid is referred to as a "target sequence" in the present invention.

The method of the present invention is characterized in that the hybridization between a probe nucleic acid and a sample nucleic acid is performed in the presence of a double-stranded DNA-binding protein.

25 The hybridization may be performed in the same manner as conventional methods for gene analysis by hybridization,

except for the use of the double-stranded DNA-binding protein. In addition, in the method of the present invention, at least one of the aforementioned probe nucleic acid or the sample nucleic acid should be DNA.

5 In the method of the present invention, it is preferable that the probe nucleic acid or the sample nucleic acid should be immobilized on a substrate, and the hybridization between the probe nucleic acid and a sample nucleic acid should be performed on the substrate.

10 In this case, it is preferable that the sample nucleic acid should be labeled to detect whether or not the hybridization has occurred. The method used for labeling is not particularly limited, and for example, it may include methods utilizing radioisotopes,  
15 fluorescent dyes, biotin and so forth.

As a substrate for hybridization, various kinds of materials, for example, membranes made of resin such as nylon membranes and nitrocellulose membranes, glass, silicon, etc are usually used. However, in the present  
20 invention, the substrate is not limited to these materials, and any substrate can be utilized so long as DNA and RNA can be immobilized on it in some way.

In this specification, the substrate on which a probe nucleic acid is immobilized is also referred to as  
25 a "DNA chip" or a "DNA array (DNA microarray or DNA macroarray)" for convenience. However, as was described earlier, this does not mean that the nucleic acid to be

immobilized must be DNA, as it may also be RNA. Those comprising membranes such as nylon membrane, nitrocellulose membrane and so forth having an immobilized probe nucleic acid are generally referred to as "DNA (macro)array", and those comprising a rigid substrate composed of glass, silicon etc. having an immobilized probe nucleic acid are sometimes referred to as "DNA chip" or "DNA (micro)array". In the present invention, all of these may be used.

10       As the method for immobilization of the probe nucleic acid or the sample nucleic acid onto the substrate, there can be mentioned the method utilizing synthesis of nucleic acids directly onto a substrate [A.C. Pease et al., Proc. Natl. Acad. Sci. USA, 91, 15 5022-5026 (1994)], the methods comprising immobilization of synthesized nucleic acids [Z. Guo et al., Nucl. Acids Res., 22, 5456-5465 (1994)] or PCR products [M. Shena et al., Proc. Natl. Acad. Sci. USA, Vol. 93, pp. 10614-10619 (1996)] and so forth onto a substrate.

20       For example, the synthesis of a probe nucleic acid or a sample nucleic acid can be performed according to a standard protocol. In the synthesis, it is preferable to synthesize them with a cycle in which trityl groups, which are usually used as a protective group for the 5' 25 end, are not removed, when the following purification method is used. The synthesized nucleic acid is preferably purified by using Poros Oligo R3 (produced by

PerSeptive Biosystems) or the like.

It is preferable to adjust the concentrations of the probe nucleic acid or the sample nucleic acid to predetermined concentrations prior to immobilization  
5 onto the substrate or hybridization thereof. For example, these nucleic acids are concentrated to dryness, and then suspended in a 0.5 M sodium hydrogencarbonate buffer (pH 8.4), TE buffer or the like, and they are quantified based on absorbance at 260 nm and adjusted to  
10 1 nmol/ $\mu$ l.

Nucleic acids can be immobilized onto the substrate, for example, by bonding amino groups of amino-modified oligonucleotides to a nylon membrane having anionic carboxyl groups on its surface at a high  
15 density through amide bonds, as described below.

As for the substrate, nitrocellulose membranes, glass and so forth can also be utilized. It is particularly preferable to utilize glass when the sample DNA will be labeled with a fluorescent substance.

20 As the method for labeling and detection of nucleic acid, the RI method utilizing [ $\gamma$ - $^{32}$ P] ATP was used in the example described later, but either the fluorescent method or the biotin-avidin method may be used as well. Examples of the fluorescent substance  
25 include Cy3, Cy5, FITC (fluorescein isothiocyanate) and so forth. As for the labeling method, in addition to the 5' end labeling method, the random primer labeling

method, the nick translation method and so forth can be utilized as well.

Methods of the aforementioned synthesis of nucleic acid, hybridization, the labeling of nucleic acids and so forth are described in references well known to those skilled in the art, for example, Maniatis, T. et al., "Molecular Cloning, A Laboratory Manual, Second Edition", Cold Spring Harbor Laboratory Press (1989)) and so forth.

Although various kinds of double-stranded DNA-binding proteins are known as double-stranded DNA-binding proteins, it is preferable to utilize a thermostable double-stranded DNA-binding protein which is derived from a hyperthermophilic bacterium (bacterium which can grow at temperatures of 90°C or higher). Hyperthermophilic bacteria are assumed to have some double-stranded DNA-binding protein because, when double-stranded DNA is separated from genome of hyperthermophilic bacteria and heated to 100°C, the DNA becomes single-stranded DNA.

Further, among hyperthermophilic bacteria, archaebacteria such as *Methanobacterium*, *Methanococcus*, *Archaeoglobus*, *Pyrococcus* and so forth contain double-stranded DNA-binding proteins such as histone-like proteins and HU protein.

Other than those, as double-stranded DNA-binding proteins derived from archaebacteria, there have been known proteins derived from *Sulfolobus* bacteria such as

Sso7d protein derived from *Sulfolobus solfataricus* [A. Guagliardi et al., J. Mol. Biol., Vol. 267, p.841-848 (1997)] ,Sac7d and Sac7e proteins derived from *Sulfolobus acidocaldarius* [J.G. McAfee et al., Biochemistry, Vol. 34, p. 10063-10077 (1995)], Sac7a and Sac7b proteins derived from *Sulfolobus acidocaldarius* [Kimura, M. et al., FEBS Letts., Vol. 176, p. 176-178(1984); Choli, T. et al., J. Biol. Chem., Vol.263, p.7087-7093(1988)] and so forth. However, it is more preferable to utilize the Sso7d protein derived from *Sulfolobus solfataricus*.

As shown in SEQ ID NO: 9, the Sso7d protein is a protein consisting of 63 amino acids [H. Baumann et al., Nature Structural Biology, Vol. 1, p. 808-819 (1994)]. In the present invention, a protein consisting of the 63 amino acids or a protein having a homology thereto of 75% or more in amino acid sequence are particularly preferably used as the double-stranded DNA-binding protein.

In the present invention, analysis of the homology is carried out by using the Lipman-Pearson method.

As software for analysis, "Genetyx" produced by Software Development Co., LTD. is used.

The result of analysis by using the method is shown as follows.

Sso7d vs. Sac7a : 84.5%

Sso7d vs. Sac7b : 87.5%

Sso7d vs. Sac7d : 84.5%

Sso7d vs. Sac7e : 83.3%

These double-stranded DNA-binding proteins have a  
5 function of stabilizing the double strands  
complementarily hybridized, and it is considered that  
they can maintain the double strands acceleratedly  
hybridized at high temperatures as it is without causing  
re-dissociation thereof. It is considered that a  
10 protein having such a function can be utilized in the  
present invention like the specifically aforementioned  
protein. Any special conditions are not particularly  
required for the expression of the aforementioned  
function of the double-stranded DNA-binding protein, and  
15 ordinary hybridization conditions with the presence of  
dithiothreitol(DTT), magnesium ions and so forth are  
sufficient.

Further, the double-stranded DNA-binding protein  
is preferably purified. The purification method may be  
20 a conventional purification method for proteins, and for  
example, purification can be performed by the method of  
Bauman (Nature Structural Biology, Vol. 1, p.808-819  
(1994)) and so forth.

Specifically, the purification of the Sso7d  
25 protein, for example, can be performed by the following  
procedure. First, the *Sulfolobus solfataricus* strain  
DSM 1618 (strain IFO 15331) is cultured, and then the

obtained bacterial cells are disrupted by a French press or the like and centrifuged. The obtained supernatant fraction is fractionated by using a MonoQ (produced by Pharmacia) column, and the target fraction is

5 concentrated. The fraction is fractionated by using a Superose 6 column, and the target fraction is dialyzed and fractionated by using a MonoS column to perform the purification.

10 An example of purification of the Sso7d protein obtained from *Sulfolobus solfataricus* is shown in the working examples mentioned hereinafter. It can also be purified from bacterial cells such as *Escherichia coli* cells in which the protein is highly produced thanks to a Sso7d protein gene introduced by a gene recombination  
15 technique.

Sub  
C1  
20 The hybridization can be performed in the same manner as the usual nucleic acid hybridization, except that the hybridization is performed in the presence of a double-stranded DNA-binding protein. Specifically the hybridization can be performed as follows. First reducing agents such as dithiothreitol (DTT) and 2-mercaptoethanol, bovine serum albumin (BSA) and skim milk which prevent protein from non-specific binding to vessel and stabilize protein are added, and further protein  
25 accessory factors such as magnesium chloride ( $MgCl_2$ ), salt-condensation regulators such as sodium chloride (NaCl) and potassium chloride (KCl) and so forth



Sub  
C1

are added as required into buffer such as Tris buffer, phosphate buffer, citric acid byffer, TES buffer, HEPES buffer or the like. The double-stranded DNA biding protein is then added to the solution. In this

5 hybridization solution, the aforementioned oligonucleotide-immobilized nylon membrane (DNA (macro)array) and a labeled sample DNA are hybridized preferably for 1-20 minutes within a range of 40-120°C. When Sso7d is used as the double-staranded DNA binding

10 protein, to a Tris buffer, 0.1-100 mM of DTT, 0.1-100 mM of MgCl<sub>2</sub> and 1-100 µg/µl of BSA (all of the concentrations are final concentrations) are added, and the Sso7d protein is added to the solution within a range of 0.001-10%. In this hybridization solution, the

15 oligonucleotide-immobilized nylon membrane (DNA (macro)array) and a labeled sample DNA are hybridized preferably for 1-15 minutes within a range of 40-70°C.

Sub  
C2

After the hybridization reaction, the membrane was washed with buffer such as Tris buffer, phosphate buffer,

20 citric acid byffer, TES buffer, HEPES buffer or the like for 1-10 minutes within a range of 10-50°C, and then dried. In this case, adding a small amount of surfactant such as sodium dodecyl sulfate (SDS) is preferable because it can prevent from non-specific

25 binding and step down backgraound. When Sso7d is used as the double-staranded DNA binding protein, after the hybridization reaction, it is preferable that the

membrane should be washed with citric acid buffer such as SSC, more preferably buffer added with SDS within a range of 0.001-0.05% as required to SSC, for 3-10 minutes within a range of 10-40 °C, and then air-dried.

- 5 As for the hybridization signal, radiation dose or the like of each dot on the dried nylon membrane can be measured by autoradiography etc. to calculate the hybridization strength.

- Detection of hybridization can be measured by a method suitable for each of various labeling methods. In the gene expression monitoring method, it is preferable to use fluorescent labeling because this method enables simultaneous detection of the expression strength for a plurality of sample nucleic acids by labeling them with a plurality of fluorescent dyes each having a different detection wavelength.

- Examples of the application of the nucleotide sequence analysis according to the present invention include: DNA nucleotide sequence determination [Genomics, Vol. 4, p. 114-128 (1989)], diagnosis of infectious and genetic diseases etc. [J.G. Hacia et al., Nature Genetics, 14, 441-447 (1996)], mapping of giant genome DNA [BioTechniques, vol.17, No.2, p.328-336 (1994)], single-nucleotide polymorphisms (SNPs) [Wang et al. Science, Vol. 280, p. 1077-1082 (1998)], the amplification of genes or analysis of deleted regions by

CGH (comparative genomic hybridization) method [D. Pinkel et al. Nat. Genet. Vol. 20, p. 207-211 (1998)], gene expression monitoring [M. Shena et al., Proc. Natl. Acad. Sci. USA, Vol. 93, pp.10614-10619 (1996)] and so  
5 forth.

As a diagnostic method for infectious diseases, there can be mentioned, for example, a method of detecting presence of a causative factor by extracting DNA from subject's blood etc., preparing a DNA probe  
10 based on a sequence specific to each of various pathogens [A. Troesch et al., J. Clin. Microbiol., Vol.37, p.49-55(1999)], and performing hybridization reaction according to the present invention to perform gene analysis for the extracted DNA.

15 As a diagnostic method for genetic diseases, there can be mentioned, for example, a method of detecting presence or absence of a mutation in a gene by preparing an oligonucleotide based on a sequence specific to a causative gene of genetic disease [M. J. Kozal et al.,  
20 Nature Medicine, Vol.2, p.753-759(1996)], and performing hybridization of the oligonucleotide with chromosomal DNA obtained from a subject according to the present invention to perform gene analysis.

The giant genome DNA mapping is an essential  
25 technique for the genome DNA analysis project and so forth. By performing hybridization of many DNA probes prepared by the method of the present invention with

genes of a genome bank, the location of each clone on the genome can be determined.

Sub  
C3  
5 Further, a test kit for using to performe the  
aforementioned gene analysis can be prepared by using a  
double-stranded DNA-binding protein. Such a test kit is  
constituted by components similar to those of ordinary  
test kits for gene analysis utilizing hybridization  
except for the use of the double-stranded DNA-binding  
protein. That is, the test kit of the present invention  
10 comprises at least a double-stranded DNA-binding protein  
and, as optional components, washing solution, diluent,  
hybridization solution and so forth.

#### EXAMPLES

15

The present invention will be explained more specifically with reference to the following examples. However, the present invention is no way limited by these examples.

20

##### Example 1

##### (A) Synthesis of probe DNA and sample DNA

The oligonucleotides shown in Table 1 were synthesized by using a DNA synthesizer (apparatus name: Expedite 8909) manufactured by PerSeptive Biosystems.

25

Immobilization of the oligonucleotides may be facilitated by modifying their 5' or 3' ends with 5'

amino-modified C6 (produced by Glen Research) or the like. The oligonucleotide numbers used below corresponds to SEQ ID NOS in SEQUENCE LISTING.

The oligonucleotide (3) is a sample DNA to be analyzed, the oligonucleotide (1) is a DNA probe completely complementary to the oligonucleotide (3) (completely matched), and the oligonucleotide (2) is a DNA probe complementary to the oligonucleotide (3), but containing one nucleotide mismatch in the internal region.

Table 1

No.	Nucleotide sequence	Note
15	(1) 5'-XATGTA <sup>SEQ ID No. 1</sup> ACTCGCCTT-3'	Completely matched probe
	(2) 5'-XATGTA <sup>SEQ ID No. 2</sup> ACCCGCCTT-3'	One base-mismatched probe
	(3) 5'-CCAACGATCAAGGCGAGTTACATGATCC-3' <sup>SEQ ID No. 3</sup>	Sample DNA

In the table, X represents 5' amino-modified C6.

The synthesis of the aforementioned oligonucleotides and the sample DNA were performed in accordance with a standard protocol (Nucleic Acid Synthesis System User's Guide, PerSeptive Biosystems) by using a cycle in which the trityl groups were not removed, which were the protective groups of the 5' ends. The synthesized DNAs were purified by using Poros Oligo R3 (produced by PerSeptive Biosystems).



10 minutes, lightly rinsed with deionized water, and air-dried.

(C) Labeling of sample DNA

5           As for the labeling of the sample DNA, the 5' end was radioactively-labeled with [ $\gamma$ - $^{32}\text{P}$ ] ATP. The reaction was performed by using a DNA 5' end labeling kit (MEGALABEL, produced by Takara Shuzo).

10   (D) Purification of Sso7d protein

          Cultivation was performed at 75°C under an aerobic condition in a medium (1 g of yeast extract, 2.3 g of  $(\text{NH}_4)_2\text{SO}_4$ , 0.08 g of  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.25 g of  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.28 g of  $\text{KH}_2\text{PO}_4$ , 0.3 g of  $\text{Na}_2\text{SiO}_3 \cdot 9\text{H}_2\text{O}$ , 0.03 g of  
15    $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ , 0.02 g of  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$  in 1 L of distilled water, adjusted to pH 3.0 with sulfuric acid) added with 10 g/L of saccharose, using a membrane fermenter. The cells were given with heat shock at 88°C for 90 minutes, and harvested by centrifugation.

20           The cells (100 g) were dissolved in Buffer A (10 mM Tris buffer, pH 8.8, 20 mM NaCl, 10% glycerol), and then disrupted with a French press. The disrupted cell suspension was centrifuged, and the obtained supernatant was dialyzed against Buffer A. Then, the dialysate was  
25   applied to a MonoQ column (produced by Pharmacia) and equilibrated with Buffer A. The Sso7d protein was obtained in the outflow fraction. This fraction was

concentrated with Amicon, applied to a Superose 6 column and equilibrated with Buffer B (30 mM Tris-HCl buffer, pH 7.4, 200 mM NaCl). The fractions containing Sso7d were collected and dialyzed against Buffer C (50 mM potassium phosphate, pH 6.0, 50 mM NaCl), and then applied to a MonoS column (produced by Pharmacia). The column was equilibrated with Buffer C and eluted with a gradient using Buffer D (50 mM potassium phosphate, pH 8.0, 1 M NaCl) to obtain the Sso7d protein at a 25% concentration of Buffer D.

(E) Hybridization reaction

(a) Hybridization strength

70 µg/ml of Sso7d protein was added to Buffer E (20 mM Tris buffer, pH 7.5, 2 mM DTT, 5 mM MgCl<sub>2</sub>, 10 µg/µl of BSA), and the aforementioned oligonucleotide-immobilized nylon membrane and the radioactively-labeled sample DNA were allowed to hybridize at 60°C for 3 minutes in the buffer. As a control, hybridization reaction was also performed in the same manner without adding the Sso7d protein.

After the hybridization reaction, the membrane was washed with 1 x SSC/0.03% SDS buffer at 30°C for 5 minutes and air-dried. The hybridization signal was evaluated by measuring radiation dose of each dot on the air-dried nylon membrane by autoradiography to calculate hybridization signal strength. The results are shown in



Table 2. The hybridization signal strength is represented with relative values based on the hybridization signal strength of the probe oligonucleotide (1) when Sso7d was added, which is taken as 100.

Table 2

No.	Nucleotide sequence	Addition of Sso7d	Hybridization signal strength	
10				
	(1) 5'-XATGTA <sup>SEA ID No. 1</sup> ACTCGCCTT-3'	Completely matched	Not added	<5
	(2) 5'-XATGTA <sup>SEA ID No. 2</sup> ACCCGCCTT-3'	Mismatched	Not added	<5
15	(1) 5'-XATGTA <sup>SEA ID No. 1</sup> ACTCGCCTT-3'	Completely matched	Added	100
	(2) 5'-XATGTA <sup>SEA ID No. 2</sup> ACCCGCCTT-3'	Mismatched	Added	10

In the table, X represents 5' amino-modified C6.

In the hybridization systems not added with Sso7d, hybridization between the probe DNA immobilized on the substrate (on the DNA array) and the sample DNA was not substantially detected for both of the completely matched probe (1) and the mismatched probe (2).

On the other hand, when the Sso7d protein was added to the hybridization systems, only the completely matched probe hybridized with the sample DNA on the DNA array. In comparison with the hybridization signal

strength of the completely matched probe, the hybridization signal strength of the mismatch probe was considerably weaker, and these were clearly distinguishable.

5

(b) Change of hybridization signal strength with change of Sso7d protein concentration

The Sso7d protein was added to Buffer E at a concentration of 70  $\mu\text{g/ml}$ , 50  $\mu\text{g/ml}$ , 25  $\mu\text{g/ml}$  or 10  $\mu\text{g/ml}$ , and the nylon membrane on which the completely matched probe in Table 1, the probe oligonucleotide (1), was immobilized, and the radioactively-labeled sample DNA were hybridized at 60°C for 3 minutes in each solution.

After the hybridization reaction, the membrane was washed with a solution of 1 x SSC/0.03% SDS buffer at 30°C for 5 minutes, and then air-dried. The hybridization signal was evaluated by measuring radiation dose of each dot on the air-dried nylon membrane by autoradiography to calculate hybridization signal strength. The results are shown in Table 3.

In the table, the hybridization signal strength is represented with relative values based on the hybridization signal strength obtained when Sso7d was added at a concentration of 70  $\mu\text{g/ml}$ , which is taken as 100.

Table 3

No. Addition of Sso7d		Hybridization signal Strength
5	(1) 70 $\mu\text{g/ml}$	100
	(2) 50 $\mu\text{g/ml}$	100
	(3) 25 $\mu\text{g/ml}$	40
	(4) 10 $\mu\text{g/ml}$	10

The completely matched probe strongly hybridized to the sample DNA on the DNA array at a Sso7d protein concentration of 50  $\mu\text{g/ml}$  or more. However, at the concentrations of 25  $\mu\text{g/ml}$  and 10  $\mu\text{g/ml}$ , the hybridization signal strength was weak. These results show that the hybridization signal sensitivity was increased by addition of Sso7d.

(c) Change of hybridization signal strength with hybridization reaction time

A nylon membrane on which the completely matched probe, the oligonucleotide (1), was immobilized, and a radioactively-labeled sample DNA were allowed to hybridize at 60°C under the same condition as the above (E) (a), except that the hybridization reaction time was changed. The hybridization signal strength was measured at 3, 10 and 30 minutes after the start of the hybridization.

The hybridization signal strength did not vary

with the reaction time. The hybridization had been almost fully completed within 3 minutes of reaction time, thus showing that the hybridization reaction speed was increased with the addition of the Sso7d protein.

5

## Example 2

### (A) Synthesis of probe DNA and sample DNA

The oligonucleotides shown in Table 4 were synthesized by using a DNA synthesizer (apparatus name: Expedite 8909) manufactured by PerSeptive Biosystems. Immobilization of the oligonucleotides may be facilitated by modifying their 5' or 3' ends with 3' amino-modified C6 (produced by Glen Research) or the like. As for the end of the sample DNA, its 5' end may be labeled with fluorescence using Cy5 Amidite (produced by Amersham Pharmacia Biotech). The oligonucleotide numbers used below corresponds to SEQ ID NOS in SEQUENCE LISTING.

The oligonucleotide (8) is a sample DNA to be analyzed, the oligonucleotides (4) and (6) are DNA probes completely complementary to the oligonucleotide (8) (completely matched), and the oligonucleotides (5) and (7) are DNA probes complementary to the oligonucleotide (8), but containing one nucleotide mismatch in the intermediary region.

Table 4

No.	Nucleotide sequences	Note
a	(4) 5'-ATCGCCCGGACTCX-3' <sup>SEA ID No. 4</sup>	Completely matched probe
a	5 (5) 5'-ATCGCCTGGACTCX-3' <sup>SEA ID No. 5</sup>	One base mismatched probe
a	(6) 5'-iiiiTCGCCCGGACTiiiiX-3' <sup>SEA ID No. 6</sup>	Completely matched probe
a	(7) 5'-iiiiTCGCCTGGACTiiiiX-3' <sup>SEA ID No. 7</sup>	One base mismatched probe
a	(8) 5'-Cy5AGTCTCGGAGTCCGGGCGATGGCCAC-3' <sup>SEA ID No. 8</sup>	Sample DNA

In the table, i represents inosine residue, X represents  
 10 3' amino-modified C6, and Cy5 represents a fluorescent  
 label by Cy5.

The syntheses of the aforementioned  
 oligonucleotides and the sample DNA were performed in  
 15 accordance with a standard protocol (Nucleic Acid  
 Synthesis System User's Guide, PerSeptive Biosystems) by  
 using a cycle in which the trityl groups were not  
 removed, which were the protective groups of the 5' ends.  
 The synthesized DNA were purified by using Poros Oligo  
 20 R3 (produced by PerSeptive Biosystems).

The oligonucleotides (4) to (8) and the sample DNA  
 were concentrated to dryness, and then suspended in a  
 0.5 M sodium hydrogencarbonate buffer (pH 8.4) for (4)  
 to (7), or TE buffer for (8), and they are quantified  
 25 based on absorbance at 260 nm and adjusted to 100  
 pmol/ $\mu$ l.

(B) Immobilization of probe DNA (preparation of DNA microarray or DNA chip)

Immobilization of the oligonucleotides was performed as follows. First, solutions of the oligonucleotide probe nucleic acids (4) to (7) were each spotted onto a Silylated Slide of TeleChem (slide glass having aldehyde groups on its surface) using a GTMASS Stamping apparatus produced by Nippon Laser & Electronics Lab. Then, according to the protocol of TeleChem, the nucleic acid was bonded on the slide glass through terminal covalent bonds (a Schiff base was formed by amino group and aldehyde group) as follows. First, the slide glass was mounted on a slide rack, and washed twice with 0.2% SDS at 25°C for 2 minutes with sufficient stirring in a beaker. Then, it was washed twice with sterilized water at 25°C for 2 minutes with sufficient stirring, and further treated with sterilized water at 98°C for 2 minutes. The slide glass was air-dried at room temperature for 5 minutes, then transferred into a sodium hydrogenborate solution [prepared by dissolving 1 g of  $\text{NaBH}_4$  in 300 ml of PBS buffer (prepared by dissolving 8 g of sodium chloride, 0.2 g of potassium chloride, 1.44 g of disodium hydrogenphosphate, and 0.24 g of potassium dihydrogenphosphate in deionized water, adjusting it to pH 7.4 with hydrochloric acid, and filling it up to 1000 ml) and 100 ml of ethanol], treated in the solution at

25°C for 5 minutes, washed three times with 0.2% SDS at 25°C (room temperature) for 1 minute, finally washed with sterilized water at 25°C for 1 minute, and air-dried.

5

(C) Hybridization reaction

70 µg/ml of the Sso7d protein and 100 pmol/ml of the sample DNA (8) were added to Buffer E (20 mM Tris buffer, pH 7.5, 2 mM DTT, 5 mM MgCl<sub>2</sub>, 10 µg/µl of BSA), and the aforementioned oligonucleotide-immobilized slide glass (DNA microarray) and the fluorescently labeled sample DNA were allowed to hybridize at 60°C for 6 minutes in the solution. As a control, hybridization reaction was also performed in the same manner without adding the Sso7d protein.

After the hybridization reaction, the membrane was washed with 1 x SSC/0.03% SDS buffer at 25°C for 5 minutes, rinsed with 0.2 x SSC, and then further rinsed with 0.5 x SSC. After the washing solution was removed by centrifugation, the slide glass was air-dried.

The hybridization signal was evaluated by measuring amount of fluorescent dye of each spot on the air-dried slide glass by using ScanArray 3000 produced by Genaral Scanning to calculate the hybridization signal strength. The results are shown in Table 5. The results are represented with relative values based on the hybridization signal strength of the probe DNA (4)

Sub  
C5

dition of Sso7

Sub  
C5

[illegible]





In the hybridization systems not added with Sso7d, hybridization between the probe DNA immobilized on the substrate (on the DNA array) and the sample DNA was not substantially detected under the aforementioned  
5 conditions.

On the other hand, when the Sso7d protein was added to the hybridization systems, only the completely matched probes hybridized with the sample DNA on the DNA array. In comparison with the hybridization signal  
10 strength of the completely matched probes, the hybridization signal strength of the mismatched probes was considerably weaker, and these were clearly distinguishable.

15 According to the present invention, gene analysis by hybridization utilizing a probe nucleic acid can be quickly performed with high precision and high sensitivity.

20 Having thus described the present invention, it will be obvious that several aspects of the invention may be modified in various ways. Such variations are not to be regarded as departures from the spirit and scope of the invention, and all such modifications would  
25 be obvious to those skilled in the arts, and are intended to be included within the scope of the following claims.

## <120> Method for Gene Analysis

<151> 1999-05-25

<210> 1

**<212> DNA**

<220>

<220>

<222> (1)

<223> n=5' amino-modified C6

natgtaactc gcott

<211> 15

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence : probe for hybridization

 $\langle 220 \rangle$ 

```
<221> misc_feature
```

/ <222> (1)

15

[illegible]

Sub  
a

<223> n=5' amino-modified C6

<400> 2

natgtaaccc gcctt

15

<210> 3

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence : sample for hybridization

<400> 3

ccaacgatca aggcgagtta catgatcc

28

<210> 4

<211> 14

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence : probe for hybridization

<220>

<221> misc/feature

<222> (14)

<223> n=~~β~~, amino-modified C6

<400> 4

atcgccgga ctcn

14

~~<210>~~ 5

<21/1> 14

<212> DNA

<213> Artificial Sequence

 $\langle 220 \rangle$ 

<223> Description of Artificial Sequence : probe for

[illegible]

Sub  
a

Sub  
A

```
<221> misc_feature
```

<223> n=3' amino-modified C6

atcgcctgga ctcn

14

<211> 20

<213> Artificial Sequence

<223> Description of Artificial Sequence : probe for hybridization

```
<221> misc_feature
```

$\langle 223 \rangle \quad n=i$

```
<221> misc_feature
```

<223> n=3' amino-modified C6

nnnntcgccc ggactnnnnn

20

<211> 20

<213> Artificial Sequence

<223> Description of Artificial Sequence : probe for hybridization

Sub  
a

<222> (1)-(4), (16)-(19)

<220>

<222> (20)

<400> 7

20

<211> 26

<213> Artificial Sequence

<223> Description of Artificial Sequence : sample for hybridization

agtctggag tccggg~~g~~cgat ggccac

26

<211> 63

<212> PRT

<213> Sulfolobus Solfataricus

<301> Herbert Baumann, Stefan Knapp, Thomas Lundback, Rudolf  
Ladenstein and Torleif Hard

<302> Solution structure and DNA-binding properties of a thermostable protein from the archaeon *Sulfolobus Solfataricus*

~~303~~ Nature structural biology

/<304> 1

/ <305> 11

<306> 808-819

<307> 1994-11-1

<400> 9

Ala Thr Val Lys Phe Lys Tyr Lys Gly Glu Glu Lys Gln Val Asp Ile  
1 5 10 15  
Ser Lys Ile Lys Lys Val Trp Arg Val Gly Lys Met Ile Ser Phe Thr  
20 25 30  
Tyr Asp Glu Gly Gly Gly Lys Thr Gly Arg Gly Ala Val Ser Glu Lys  
35 40 45  
Asp Ala Pro Lys Glu Leu Leu Gln Met Leu Glu Lys Gln Lys Lys  
50 55 60

Sub  
a!

[illegible]